

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

This Page Blank (uspto)

für A 1
①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

Offenlegungsschrift
⑪ DE 3938627 A1

⑤1 Int. Cl. 5:
G06K 9/64
B 07 C 3/14
// G10L 7/10

②1 Aktenzeichen: P 39 38 627.9
②2 Anmeldetag: 21. 11. 89
④3 Offenlegungstag: 31. 5. 90

DE 3938627 A1

③0 Unionspriorität: ③2 ③3 ③1
30.11.88 JP 302692/88

⑦1 Anmelder:
Kabushiki Kaisha Toshiba, Kawasaki, Kanagawa, JP

⑦4 Vertreter:
Kramer, R., Dipl.-Ing.; Weser, W., Dipl.-Phys.
Dr.rer.nat.; Hoffmann, E., Dipl.-Ing., 8000 München;
Blumbach, P., Dipl.-Ing.; Zwirner, G., Dipl.-Ing.
Dipl.-Wirtsch.-Ing., Pat.-Anwälte, 6200 Wiesbaden

⑦2 Erfinder:
Yamagishi, Kaiko, Tokio/Tokyo, JP

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤4 Mustererkennungsvorrichtung

Eine Mustererkennungsvorrichtung liest das Eingabemuster und erkennt das Eingabemuster, indem sie die Auswertepunkte berechnet, welche den Grad der Ähnlichkeit zwischen dem gelesenen Eingabemuster und einem Standardmuster in mehreren Ebenen ausdrücken.

DE 3938627 A1

Die Erfindung bezieht sich auf eine Mustererkennungsvorrichtung, die zum Beispiel in einem optischen Zeichenlesegerät oder einer Stimmerkennungsvorrichtung eingesetzt wird.

In den letzten Jahren wurde auf dem Gebiet der Mustererkennungstechnik von Zeichen oder Stimmen oder ähnlichem häufig ein Prozeß angewandt, bei dem die Erkennungsergebnisse des zuletzt eingegebenen Musters dadurch erhalten wurden, daß man in einer Datenbank, zum Beispiel einem Wortverzeichnis, auf der Grundlage der aus den Mustererkennungsergebnissen erhaltenen Daten recherchierte.

Bei einer solchen Mustererkennungsvorrichtung wird z. B. zuerst das Zeichenmusterbild auf der Oberfläche einer Postkarte *P* optisch gelesen, die gelesenen Daten werden einer A/D-Umsetzung unterzogen, dann wird jedes Zeichen dadurch erkannt, daß Auswertungspunkte berechnet werden, die das Maß an Ähnlichkeit zwischen dem Bildzeichenmuster und dem Normmuster ausdrückt, und zwar mit Hilfe eines Ähnlichkeitsmaß-Verfahrens, wie es in der US-PS 36 88 267 beschrieben ist, um so eine Zeichen-Kandidatengruppe von einem ersten bis zu einem vierten Kandidaten zu erhalten, wobei weiterhin als Ergebnis des Erkennungsvorgangs Wortbegrenzungsinformation erhalten wird.

Dann wird für die Wort-Recherche auf der Grundlage des ersten Kandidaten für jedes Zeichen und von Wort-Begrenzungsinformation ein Schlüsselwort für die Wortrecherche generiert, und dieses Schlüsselwort einerseits und die Wörter in dem Wortverzeichnis andererseits werden beispielsweise für jeweils einen drei Zeichen umfassenden Zug verglichen, um sämtliche Wörter aus dem Wortverzeichnis zu extrahieren, bei denen Übereinstimmung mit dem Zeichenzug herrscht.

Danach wird jedes als Ergebnis der Recherche erhaltene Wort in einen Wort-ID-Code umgesetzt, und es wird eine Recherche in einer Adreß-Datenbank durchgeführt, wobei der ID-Code als Schlüsselwort dient, um so als letztendliche Erkennungsergebnisse den Adreßcode zu erhalten, der der auf der Postkarte *P* notierten Adresse entspricht, d. h. der dem Postleitzahl-Code entspricht.

Allerdings sind bei der herkömmlichen Vorrichtung folgende Dinge zu berücksichtigen:

Wenn die auf der Postkarte *P* notierten Zeichen insgesamt sehr hell sind oder schwach sind, so war es bislang nicht möglich, eine ausreichende Mustererkennung bei üblichen Bedingungen durchzuführen, so daß es Schwierigkeiten bei dem getreuen Herausziehen der Zeichen-Kandidaten gab, die Merkmale der Eingabezeichen-Muster aufwiesen. Ferner gab es das Problem, daß Lesefehler bei der Zeichenerkennung einen ungünstigen Einfluß auf den Prozeß insgesamt hatten, so daß dadurch die Zuverlässigkeit stark herabgesetzt wurde.

Weiterhin ergab sich folgendes Problem: Da sämtliche Wörter, die mit dem Schlüsselwort übereinstimmten, aus dem Wortverzeichnis durch einen 3-Zeichen-Zug bei der Durchführung der Wortrecherche herausgesucht wurden, brauchte es beträchtliche Zeit, um die Adreß-Datenbank mit den so gefundenen Werten als Schlüsselwort zu recherchieren, falls eine große Menge von extrahierten Wort-Kandidaten vorhanden war.

Bei der bekannten Vorrichtung wird die Gesamtsumme von jedem Wort aus dem Auswertungspunkt bei jedem Wort, das aus den Ergebnissen der Zeichenerkennung erhalten wurde, ermittelt und geteilt durch die Anzahl der Wörter, um so das Wort-Auswerteverhältnis zu erhalten. Die Wörter mit dem höchsten Wort-Auswerteverhältnis unter jedem der Wort-Kandidaten werden in Reihenfolge umgesetzt in die ID-Codes, und anschließend erfolgt unter ihrer Verwendung die Datenrecherche in der Adreß-Datenbank.

Allerdings ergibt sich durch das Wort-Auswerteverhältnis bei diesem Prozeß folgendes Problem: Wenn zum Beispiel als Ergebnis der Wortrecherche mit dem Schlüsselwort [JOHN] das Wort [JOHNS] erhalten wird, so ergibt sich unter der Annahme, daß die Gesamtsumme der Wort-Auswertepunkte für dieses Wort $98 \times 4 = 260$ HEX beträgt, das Wort-Auswerteverhältnis 79 HEX, was man durch Teilung durch die Wortzeichenzahl 5 erhält.

Wenn hingegen das mit dem Schlüsselwort [JOHNS] erhaltene Wort [JOHN] ist, so ist es in hohem Maße wünschenswert, daß das Verhältnis genauso groß ist wie das Wort-Auswerteverhältnis. Allerdings ist das Verhältnis (98 HEX) größer als das genannte Verhältnis, da man das Gesamtergebnis durch 4 teilt.

Ferner bestand die Möglichkeit, daß dieses Wort-Auswerteverhältnis übereinstimmte mit dem Verhältnis, was man im Fall des Schlüsselworts [JOHN] erhielt, oder das für das Wort [JOHNS] berechnete Auswerteverhältnis, das gleich dem Schlüsselwort [JOHNS] ist, war niedriger als das für das Wort [JOHN] berechnete, und zwar wegen der größeren Zeichenzahl.

Bei der herkömmlichen Vorrichtung bestand gemäß den obigen Ausführungen das Problem darin, daß das Wort-Auswerteverhältnis mit höherer Zuverlässigkeit nicht stabil ermittelt werden konnte.

Weiterhin wurde in der Vergangenheit die Wortrecherche stets für sämtliche Wörter des Wortverzeichnisses anhand einer binären Suche durchgeführt, da die Zeichencodenzüge, aus denen sich die Wörter zusammensetzen, innerhalb des Wortverzeichnisses mit Wort-ID-Codes für die Datenbankrecherche registriert sind, so daß sie im ASCII-Code in ansteigender Reihenfolge angeordnet sind.

Deshalb ergab sich ein weiteres Problem dadurch, daß mit zunehmender Anzahl von Wörtern in dem Wortverzeichnis mehr Zeit für die Wortrecherche benötigt wurde.

Zusammengefaßt: Bei der herkömmlichen Mustererkennungsvorrichtung bestand ein Problem in der Genauigkeit der Mustererkennungs-Verarbeitung und der Wortrecherche oder ein Problem hinsichtlich der Effizienz der Wortrecherche, so daß es schwierig war, die Mustererkennung mit hoher Zuverlässigkeit stabil durchzuführen.

Aufgabe der vorliegenden Erfindung ist es, bei einer Mustererkennungsvorrichtung die Genauigkeit der Mustererkennung zu erhöhen und eine Mustererkennungsvorrichtung zu schaffen, bei der die Zuverlässigkeit spürbar erhöht ist, indem die Wortrecherchen-Effizienz verbessert ist.

Die Lösung dieser Aufgabe ist in den Patentansprüchen angegeben. Vorteilhafte Weiterbildungen ergeben sich aus den Unteransprüchen.

Eine Optimalmuster-Auswahleinrichtung wählt das optimale Muster aus auf der Grundlage der Auswertepunkte auf jeder Ebene, die durch die Mustererkennungseinrichtung berechnet wurden.

Die Mustererkennungsvorrichtung gemäß der Erfindung enthält eine Mustererkennungseinrichtung, die die Auswertepunkte berechnet, welche das Maß der Ähnlichkeit zwischen dem Eingabemuster von der Musterleseeinrichtung und dem Norm- oder Standardmuster in jedem von mehreren Ebenen ausdrückt, um das Eingabemuster zu erkennen. Die Optimalmuster-Auswahleinrichtung wählt auf der Grundlage der Auswertepunkte in jeder Ebene, berechnet durch die Mustererkennungseinrichtung, das optimale oder beste Muster aus, so es möglich ist, genaue Mustererkennungsergebnisse zu erzielen, ungeachtet der Tatsache, ob das Eingabemuster schwach oder abgeschattet ist. Die Leistungsfähigkeit der Mustererkennung wird also erhöht.

Im folgenden werden Ausführungsbeispiele der Erfindung anhand der Zeichnung näher erläutert. Es zeigt

Fig. 1 ein Blockdiagramm, welches den Aufbau eines Postleitzahl-Lesegeräts mit Hilfe einer ersten Ausführungsform der Mustererkennungsvorrichtung veranschaulicht;

Fig. 2 ein Flußdiagramm des Mustererkennungsprozesses innerhalb des Postleitzahl-Lesegeräts gemäß Fig. 1;

Fig. 3 eine Ansicht von Beispielen einer Zeichenkandidaten-Gruppe, die man durch den Zeichenerkennungsabschnitt in Fig. 1 enthält, und die Auswertepunkte;

Fig. 4 eine Ansicht von Schlüsselwörtern, die aus den Ergebnissen der Zeichenerkennung gemäß Fig. 3 generiert werden;

Fig. 5 ein Flußdiagramm der Wortrechercheverarbeitung bei einem zweiten Ausführungsbeispiel;

Fig. 6 eine Ansicht, die die Ergebnisse der Trigrammrecherche bei dieser Ausführungsform veranschaulicht;

Fig. 7 ein Flußdiagramm der Wortrechercheverarbeitung der dritten Ausführungsform;

Fig. 8 ein Flußdiagramm der Wortrechercheverarbeitung gemäß der vierten Ausführungsform; und

Fig. 9 eine Darstellung der Ergebnisse der Zeichenerkennung bei dieser Ausführungsform.

Fig. 1 ist ein Blockdiagramm, welches den Aufbau eines Postleitzahl-Lesegeräts darstellt, in Verbindung mit einer Ausführungsform der erfindungsgemäßen Mustererkennungsvorrichtung.

Wie in der Zeichnung dargestellt, besteht das Postleitzahl-Lesegerät aus einem Leseabschnitt 11, der das Zeichenmuster auf der Oberfläche einer Postkarte *P* durch optisches Abtasten der Postkarte *P*, auf der eine Adresse, eine Postleitzahl und dergleichen aufgezeichnet sind, liest, einem Zeichenerkennungsabschnitt 12, der die Zeichenerkennung durchführt, indem er die Auswertepunkte berechnet, welche das Ähnlichkeitsmaß zwischen dem Eingabe-Zeichenmuster vom Leseabschnitt 11 und dem Standardmuster an mehreren Schnittebenen ausdrücken, die zuvor in bezug auf die Färbung des Zeichens eingestellt wurden, einem Wortverzeichnis 13, in welchem aus Zeichenzügen bestehende Wörter gespeichert sind, einem Wortrecherchenabschnitt 14, welcher aus von dem Zeichenerkennungsabschnitt 12 erhaltenen Zeichenkandidaten-Gruppen die Zeichenkandidaten mit höherem Auswertmaß für jedes Zeichen und Wortbegrenzungsinformation auswählt sowie das Schlüsselwort generiert, welches aus den Zeichenzügen besteht, indem er diese Zeichenkandidaten kombiniert, um die Wortrecherche auf der Basis des Schlüsselworts in dem Wortverzeichnis 13 durchzuführen, einer Datenbank 15, in der Codedaten von jeder Adresse gespeichert sind, und einem Datenbankrecherchenabschnitt 16, der die Recherche in der Adreß-Datenbank 15 auf der Grundlage der von dem Wortrecherchenabschnitt 14 erhaltenen Wörter durchführt. Durch diese Elemente wird der Mustererkennungsmechanismus gebildet.

Weiterhin ist die Vorrichtung nach diesem Ausführungsbeispiel mit einem Ablage- oder Stapelabschnitt 18 ausgestattet, wo die vom Leseabschnitt 11 gelesenen Postkarten nach Maßgabe des zu speichernden Adreßcodes klassifiziert werden. Ein Postkartentransferabschnitt 19 transferiert die Postkarten *P* nach dem Lesen zu dem Stapelabschnitt 18. Ein Transfersteuerabschnitt 17 steuert den Postkartentransferabschnitt 19 derart, daß dieser die Postkarten *P* auf der Grundlage des von dem Datenbankrecherchenabschnitt 16 zugeführten Adreßcodes in das zugehörige Fach transferiert. Diese Einzelteile sind Zubehörelemente des Mustererkennungsmechanismus.

Im folgenden soll anhand des in Fig. 2 dargestellten Flußdiagramms die Arbeitsweise dieser Ausführungsform erläutert werden.

Der Leseabschnitt 11 liest optisch das Zeichenmusterbild auf der Postkarte *P* und transferiert die Musterbilddaten über einen Analog-/Digital-Umsetzer zu dem Zeichenerkennungsabschnitt 12 (Schritt 201).

Dann erkennt der Zeichenerkennungsabschnitt 12 jedes Zeichen, indem er die Auswertepunkte berechnet, welche das Ähnlichkeitsmaß ausdrücken zwischen dem eingegebenen Zeichenmuster und dem Standardmuster an mehreren Schnittebenen. Das Standardmuster wurde vorab in bezug auf die Färbung der Zeichen eingegeben. Dies geschieht, um die Zeichen-Codedaten von dem ersten bis zum vierten Kandidaten als Ergebnisse der Erkennung und die Wortbegrenzungsinformation zu dem Wortrecherchenabschnitt 14 zu übertragen (Schritt 202).

Fig. 3 zeigt den ersten bis vierten Wortkandidaten und die Auswertepunkte, die jedem Kandidaten beigelegt sind, wobei es sich um die Ergebnisse der Zeichenerkennung handelt, wenn auf der Postkarte *P* der Wort-Zeichenzug [ROOM] notiert ist.

Weiterhin kennzeichnet die Mustergruppe 1 diejenigen Zeichenkandidaten, die in einer regulären Auswertebene erhalten werden, und die dazugehörigen Auswertepunkte, während die Mustergruppe 2 diejenigen Zeichenkandidaten und zugehörigen Auswertepunkte kennzeichnet, die in einer Auswertebene erhalten werden, bei der angenommen wird, daß jedes Zeichen in einer schwachen Farbtönung aufgezeichnet ist.

Nach der Zeichenerkennung extrahiert der Wortrecherchenabschnitt 14 vorzugsweise zeichenweise die Zeichenkandidaten mit höheren Auswertepunkten aus den so erhaltenen Zeichenkandidaten, um das Schlüsselwort für die Wortrecherche zu erhalten, indem auf der Grundlage der Wortbegrenzungsinformation das Wort begrenzt wird (Schritt 203). Wenn dabei als erster Kandidat gleiche Zeichenkandidaten existieren, so wird der Kandidat mit der höheren Auswertepunktezahlgelassen, um ihn bei der späteren Verarbeitung zu verwenden.

den.

Fig. 4 zeigt die Zeichenkandidaten von dem ersten bis zu dem vierten Zeichen als Ergebnis der Erkennung gemäß Fig. 3.

Wie in dieser Figur gezeigt ist, ergibt sich der Zeichenzug [ROON] aus dem jeweils ersten Kandidaten der Zeichenkandidatengruppe gemäß Fig. 3.

Anschließend bringt der Wortrecherchenabschnitt 14 dieses Schlüsselwort in Vergleich mit den in dem Wortverzeichnis 13 erhaltenen Wörtern unter Verwendung von drei Zeichen umfassenden Zeichenzügen, wobei dieser Zeichenzug als Schlüsselwort für die Wortrecherche zugrunde gelegt wird (Schritt 204). Die übereinstimmenden Wörter werden aus dem Wortverzeichnis 13 entnommen, um sie in einem speziellen Pufferspeicher zu speichern (Schritt 205).

Im allgemeinen wird diese Wortrecherche als "Trigramm-Recherche" bezeichnet.

Bezüglich des Zeichenzugs [ROON] werden die Wörter herausgefunden, die 4 Arten von 3 Zeichen umfassenden Zügen enthalten: [ROO], [OON], [ONR] und [NRO].

Wenn der Vergleich mit sämtlichen im Wortverzeichnis 13 gespeicherten Wörtern beendet ist (Schritt 206), erhält der Wortrecherchenabschnitt 14 für sämtliche im Puffer gespeicherten Wörter das Auswerteverhältnis (Schritt 207).

Dieses Wortauswerteverhältnis wird dadurch erhalten, daß die Gesamtsumme der Auswertepunkte jedes Zeichens dividiert wird durch die Zeichenzahl pro Wort.

Wenn zum Beispiel das Wort [ROOM] im Puffer gespeichert ist und das Wortauswerteverhältnis dieses Wortes berechnet wird, erhält man

$$(90 + 98 + 98 + 88)/4 = 92 \text{ HEX.}$$

Nach dem Berechnen des Wortauswerteverhältnisses sämtlicher aus dem Wortverzeichnis 13 geholter Wörter liefert der Wortrecherchenabschnitt 14 die ranghöchsten acht Wörter mit höherem Wortauswerteverhältnis als Recherchenergebnis (Schritt 208).

Die so ausgewählten Wortkandidaten werden an den Datenbankrecherchenabschnitt 16 geliefert, und in der Adreßdatenbank 15 des Recherchenabschnitts 16 wird unter Verwendung jedes Wort-ID-Codes als Schlüsselwort recherchiert, der jedem Wortkandidaten entspricht (Schritt 209).

Danach werden die Ergebnisse des Erkennungsvorgangs in der Adreßdatenbank 15 an den Transfersteuerabschnitt 17 geliefert, und anschließend wird das Stapelabschnitt-Steuersignal, welches dem Adreßcode entspricht, vom Transfersteuerabschnitt 17 an den Stapelabschnitt 18 geliefert.

Der Stapelabschnitt 18 speichert die von dem Postkartentransferabschnitt 19 kommende Postkarte auf der Grundlage des Stapelsteuersignals an den Stapelabschnitt (Schritt 210).

Demgemäß wird bei der erfindungsgemäßen Vorrichtung jedes Zeichen erkannt, indem die Auswertepunkte, ausgedrückt als Ähnlichkeitsmaß oder -grad zwischen dem Eingabe-Zeichenmuster und dem Standardmuster, auf mehreren Ebenen berechnet, die eingestellt sind in Beziehung zu der Helligkeit oder Deutlichkeit des Zeichens, um das Schlüsselwort für die Wortrecherche zu erzeugen. Dies geschieht durch bevorzugtes Sammeln von Zeichenkandidaten mit höheren Auswertepunkten aus der Zeichenkandidaten-Gruppe für jedes Zeichen. Dadurch ist es möglich, das Eingabezeichenmuster korrekt zu erkennen, ungeachtet der Stärke, mit dem die Zeichen auf der Postkarte *P* notiert sind.

Im folgenden wird eine zweite Ausführungsform erläutert.

Die Postleitzahlesevorrichtung bei dieser Ausführungsform ist mit einem Wortrecherchenabschnitt 14 ausgestattet, der eine neue Wortrecherchenfunktion aufweist gegenüber dem ersten Ausführungsbeispiel, während die anderen Bestandteile die gleichen wie in Fig. 1 sind.

Bei dieser Ausführungsform ist der Wortrecherchenabschnitt 14 derart ausgestaltet, daß er das Schlüsselwort für die Wortrecherche aus denjenigen Zeichenkandidaten erzeugt, die von dem Zeichenerkennungsabschnitt 12 erhalten werden. Das Schlüsselwort wird verglichen mit den im Wortverzeichnis 13 enthaltenen Wörtern anhand einer spezifizierten Anzahl von Zeichenzügen, um ein solches Wort zu erhalten, welches eine höhere Häufigkeit an Übereinstimmungen bei den Endergebnissen der Recherche aufweist.

Die Arbeitsweise der Ausführungsform wird anhand des in Fig. 5 dargestellte Flußdiagramms erläutert.

Zunächst liest der Leseabschnitt 11 auf optischem Wege das Zeichenmusterbild auf der Postkarte *P* und überträgt die Daten mittels A-/D-Umsetzung zu dem Zeichenerkennungsabschnitt 12 (Schritt 501).

Der Zeichenerkennungsabschnitt 12 erkennt jedes Zeichen, indem er die Auswertepunkte berechnet, die das Ähnlichkeitsmaß zwischen dem eingegebenen Zeichenmuster und dem Standardmuster ausdrücken. Dies geschieht zum Beispiel mit dem Verfahren des zusammengesetzten Ähnlichkeitsmaßes in mehreren Auswerteebenen, um die Zeichencodedaten von dem ersten bis zu dem vierten Kandidaten und die Wortbegrenzungsinformation als Erkennungsergebnis an den Wortrecherchenabschnitt 14 zu liefern (Schritt 502).

Danach erzeugt der Wortrecherchenabschnitt 14 die Schlüsselwörter für die Wortrecherche aus dem Bereich des ersten bis vierten Kandidaten der Zeichen-Kandidatengruppen und der von dem Zeichenerkennungsabschnitt 12 erhaltenen Wortbegrenzungsinformation (Schritt 503).

Nacheinander vergleicht der Wortrecherchenabschnitt 14 das Schlüsselwort als ersten Kandidaten mit den Wörtern im Wortverzeichnis 13 durch Gegenüberstellung mit jeweils einem drei Zeichen umfassenden Zeichenzug (Schritt 504), und extrahiert sämtliche übereinstimmenden Wörter aus dem Wortverzeichnis 13, um sie in dem spezifizierten Datenpuffer zu speichern (Schritt 505), während die Anzahl von Übereinstimmungen für jedes Wort mit Hilfe eines Zählpuffers gezählt wird (Schritt 506).

Die in dem Wortverzeichnis 13 für die Recherche enthaltenen Wörter werden begrenzt auf Wörter, die aus der Anzahl von Zeichen bestehen, die hinsichtlich der Anzahl der Zeichen des Schlüsselworts in einem Bereich

von ± 1 liegen.

Wenn zum Beispiel der Zeichenzug [ANGLES] als Schlüsselwort für die Wortrecherche vorliegt und dieses Schlüsselwort verglichen wird mit den Wörtern im Wortverzeichnis 13 auf der Grundlage von 3-Zeichen-Zügen, so wird das Schlüsselwort viermal dem Wort [ANGELES] oder [BNGELES] jedem 3-Zeichen-Zug [ANG], [NGL], [GLE], [LES], [ESA] und [SAN] gegenübergestellt, wie es in Fig. 6 gezeigt ist, und es erfolgt eine sechsmalige Gegenüberstellung mit dem Wort, welches dem Schlüsselwort identisch ist.

Die so gezählte Anzahl oder Häufigkeit von Übereinstimmungen zeigt das Maß von Übereinstimmung mit dem Schlüsselwort an, und jedes Wort wird auf der Grundlage des Schlüsselworts in dem Pufferspeicher gespeichert.

Nachdem der Wortrecherchenabschnitt 14 die Anzahl von Übereinstimmungen mit jedem Wort gezählt hat, indem er sämtliche Wörter innerhalb des Wortverzeichnisses 13 verglichen hat (Schritt 507), wertet er jeden Zeichenkandidaten auf der Grundlage der Anzahl von Übereinstimmungen aus (Schritt 508). Nachdem der Wortrecherchenabschnitt 14 die Wörter aus den Wortkandidaten extrahiert hat, bei denen die Anzahl von Übereinstimmungen höher ist als die spezifizierte Zahl, erfolgt eine Umwandlung dieser Wörter in Wort-ID-Codes, die dem Datenbankrecherchenabschnitt 16 zugeleitet werden (Schritt 509).

Danach recherchiert der Datenbankrecherchenabschnitt 16 die Adreßdatenbank 15, wobei jeder Wort-ID-Code als Schlüsselwort dient (Schritt 510).

Der als Ergebnis der Recherche in der Adreßdatenbank 15 gelieferte Adreßcode wird dem Transfersteuerabschnitt 17 zugeführt. Der Transfersteuerabschnitt 17 liefert entsprechend dem Adreßcode an den Stapelabschnitt 18 ein Stapelsteuersignal.

Der Stapelabschnitt 18 legt die Postkarten *P*, die von dem Postkartentransferabschnitt 19 kommen, auf der Grundlage des Stapelsteuersignals in dem entsprechenden Fach des Stapelabschnitts ab (Schritt 511).

Bei dieser zweiten Ausführungsform wird das Schlüsselwort für die Wortrecherche den Wörtern innerhalb des Wortverzeichnisses 13 gegenübergestellt durch jeden 3-Zeichen-Zug, um vorzugsweise Wortkandidaten für das letztliche Recherchenwort zu erhalten, die eine höhere Anzahl von Übereinstimmungen aufweisen, so daß es möglich ist, den Wortkandidaten mit höherem Übereinstimmungsmaß aus dem Wortverzeichnis 13 selektiver zu extrahieren, wodurch die Wirksamkeit der Datenrecherche in der Adreßdatenbank 15 wesentlich verbessert wird.

Im folgenden wird eine dritte Ausführungsform erläutert.

Diese Ausführungsform besitzt einen Wortrecherchenabschnitt 14 mit einer weiteren Wortrecherchenfunktion, die zusätzlich zu der zweiten Ausführungsform vorgesehen ist. Die anderen Abschnitte entsprechen den in Fig. 1 dargestellten Einzelheiten.

Dieser Wortrecherchenabschnitt 14 ist so aufgebaut, daß er das Schlüsselwort den Wörtern innerhalb des Wortverzeichnisses 13 durch eine spezifizierte Anzahl jedes Zeichen-Zugs gegenüberstellt und weiterhin das Schlüsselwort mit jedem Zeichen vergleicht, um bevorzugt diejenigen Wortkandidaten als letztliches recherchiertes Wort zu liefern, die die höhere Anzahl von Zeichenübereinstimmungen aufweisen.

Fig. 7 zeigt anhand eines Flußdiagramms die Wortrecherche bei diesem Ausführungsbeispiel.

Der Prozeß bis zum Schritt 709 entspricht dem Ablauf bis zum Schritt 509 in Fig. 5, so daß auf eine nochmalige Erläuterung dieser Schritte verzichtet wird.

Nachdem die Wortkandidaten mit höherer Anzahl von Übereinstimmungen aus dem Wortverzeichnis 13 anhand der Übereinstimmungsergebnisse zwischen Schlüsselwort und den Wörtern innerhalb des Wortverzeichnisses unter Zugrundelegung von 3-Zeichen-Zügen extrahiert worden sind, stellt der Wortrecherchenabschnitt 14 das Schlüsselwort mit jedem Wortkandidaten zeichenweise gegenüber, um diejenigen Wortkandidaten auszuwählen, die ein höheres Maß an Übereinstimmung aufweisen (Schritt 710).

Danach beläßt der Wortrecherchenabschnitt 14 schließlich die Wortkandidaten mit der höheren Anzahl von übereinstimmenden Zeichen bis zu der spezifizierten Anzahl als Schlüsselwort für die Datenbankrecherche, während die anderen Wortkandidaten fallengelassen werden (Schritt 711).

Wenn zum Beispiel ein Zeichenzug [SUITE] auf der Postkarte *P* notiert ist und als Schlüsselwort ein Wort mit einem falschen Zeichen erhalten wird, zum Beispiel [SULTE], so ist von der Anzahl von Übereinstimmungen das Zweifache das Maximum als Ergebnis des Vergleichs mit den Wörtern in dem Wortverzeichnis 13 bei Gegenüberstellung von 3-Zeichen-Zügen, so daß eine beträchtlich große Anzahl von Wörtern wie [SUITE], [ULTERA] und [RALTES] und dergleichen als Wortkandidaten anfallen.

Wenn diese Wortkandidaten weiterhin zeichenweise dem Schlüsselwort gegenübergestellt werden, so ergibt sich keine Übereinstimmung mit Kandidaten wie [ULTERA], [RALTES] und dergleichen, jedoch ergibt sich eine 4-Zeichen-Übereinstimmung mit dem Wort [SUITE].

In diesem Fall werden also die Wörter bis zu der festgelegten hohen Übereinstimmungszahl, einschließlich des Wortes [SUITE], von dem Wortrecherchenabschnitt 14 als Schlüsselwörter für die Datenbankrecherche geliefert.

Bei der dritten Ausführungsform der Erfindung können also die Wortkandidaten mit einem noch höheren Übereinstimmungsmaß aus den durch die Trigramm-Recherche erhaltenen Wortkandidaten als letztliches Recherchenwort ausgewählt werden, so daß überflüssige Recherchierzeit zum anschließenden Recherchieren in der Datenbank verkürzt werden kann.

Bei dieser Ausführungsform wählt die Vorrichtung den Wortkandidaten mit dem noch höheren Übereinstimmungsmaß aus den Wortkandidaten aus, die durch die Trigrammrecherche erhalten werden. Es ist jedoch zulässig, das letztliche optimale Wort dadurch auszuwählen, daß das Schlüsselwort, welches durch die Kombination der Ergebnisse der Zeichenerkennung ohne Durchführung der Trigrammrecherche erhalten wurde, direkt zeichenweise verglichen wird mit den Wörtern in dem Wortverzeichnis 13.

Im folgenden wird eine vierte Ausführungsform der Erfindung erläutert.

Diese Ausführungsform besitzt einen Wortrecherchenabschnitt 14 mit einer gegenüber der zweiten Ausführungsform neuen Wortauswerteverhältnis-Berechnungsfunktion, wohingegen die übrigen Teile der Darstellung nach Fig. 1 entsprechen.

Der Wortrecherchenabschnitt 14 dieser Ausführungsform ist derart ausgebildet, daß er die Gesamtsumme der Auswertepunkte zeichenweise für die Wortkandidaten dividiert durch die größere Zahl aus der Anzahl der Wortzeichen oder der Anzahl der Schlüsselwortzeichen, um bevorzugt denjenigen Wortkandidaten als letztliches ermitteltes Zeichen an den Datenbankrecherchenabschnitt 16 zu liefern, der das höhere Verhältnis aufweist. Anhand des Flußdiagramms (Fig. 8) soll die Berechnung des Wortauswertemaßes erläutert werden.

Zunächst wird aus dem Wortverzeichnis 13 mit Hilfe der Trigrammrecherche des Schlüsselwortes mit den Wörtern des Wortverzeichnisses durch jeden 3-Zeichen-Zug eine Mehrzahl von Wortkandidaten ausgesucht. Danach wird jedes Wortauswertemaß erhalten aus der Gesamtsumme der Auswertepunkte bei jedem Zeichen in jedem Wortkandidaten.

Das Verfahren, mit dessen Hilfe dieses Wortauswertemaß erhalten wird, hat folgenden Ablauf:

Wenn zum Beispiel der Wortkandidat [ROOM] bezüglich des Schlüsselworts [ROON] aus dem Wortverzeichnis extrahiert wird, wird, da die Zeichenzahl in beiden Wörtern gleich ist, die Gesamtsumme der Auswertepunkte jedes Zeichens in diesem Wort dividiert durch die Zeichenzahl, um sie zu dem Wortauswertemaß des Wortkandidaten zu machen.

Im vorliegenden Fall erhält man damit

$$(90 + 98 + 98 + 88)/4 = 92 \text{ HEX.}$$

Wenn weiterhin als Wortkandidat [ROOMS] extrahiert wird, erhält man das Wortauswertemaß wegen der unterschiedlichen Zeichenzahl wie folgt:

Zunächst wird der obere Zeichenabschnitt des Schlüsselworts [ROON] dem Wortkandidaten [ROOMS] gegenübergestellt, um zwischen jedem Zeichen die Auswertepunkte zu erhalten. In diesem Fall bedeutet dies:

R:R 90 HEX

O:O 98 HEX

O:O 98 HEX

N:M 88 HEX

Dann werden die Auswertepunkte nacheinander von oben nach unten addiert. In diesem Fall werden die Punkte um die Länge des kurzen Zeichens addiert. Daraus erhält man

R:R 90 HEX 90 HEX

O:O 98 HEX 128 HEX

O:O 98 HEX 1C0 HEX

N:M 88 HEX 248 HEX

Dann werden die letzten Zeichen in dem Schlüsselwort [ROON] und dem Wortkandidaten [ROOMS] gegenübergestellt, um die Auswertepunkte für jedes Zeichen zu erhalten. Man erhält:

R:O 50 HEX

O:O 98 HEX

O:M 50 HEX

N:S 50 HEX

Dann werden die Auswertepunkte der Zeichen von unten nach oben addiert, und man erhält

R:O 50 HEX 188 HEX

O:O 98 HEX 138 HEX

O:M 50 HEX A0 HEX

N:S 50 HEX 50 HEX

Dann wird das Verhältnis, welches ermittelt wurde durch Gegenüberstellung der nach beiden Verfahren ermittelten Verhältnisse, jeweils um eine Spalte oder Reihe nach hinten verschoben, und es erfolgt eine Ausrichtung mit dem jeweiligen Verhältnis, welches durch Gegenüberstellung der letzten Zeichen erhalten wurde, um die Summe zu gewinnen. Dadurch erhält man folgendes Ergebnis:

$$\begin{array}{rclcl}
 & 188 & \text{HEX} & = & 188 & \text{HEX} \\
 90 & \text{HEX} & + & 138 & \text{HEX} & = & 1\text{C}8 & \text{HEX} \\
 128 & \text{HEX} & + & \text{A}0 & \text{HEX} & = & 1\text{C}8 & \text{HEX} \\
 1\text{C}0 & \text{HEX} & + & 50 & \text{HEX} & = & 210 & \text{HEX} \\
 248 & \text{HEX} & & & & = & 248 & \text{HEX}
 \end{array}$$

Von diesen bildet der Maximalwert (248 HEX) die Gesamtsumme der Auswertepunkte.

Das Wort-Auswertemaß für diese Gesamtsumme wird berechnet, indem man die Gesamtsumme teilt durch die Anzahl von Zeichen von [ROOMS], und auf diesen Wert A HEX als Kompensationspunkt addiert.

Daher erhält man durch die Wortauswertung in diesem Fall:

$$(24 \text{ HEX} + 50 \text{ HEX})/5 + \text{A HEX} = 7\text{E HEX}.$$

Danach wird das Wortauswertemaß verglichen mit dem Auswerteverhältnis (92 HEX), welches für [ROOM] berechnet wurde, um den Wortkandidaten mit dem höheren Auswerteverhältnis als letztes Recherchenwort zu erhalten.

Die Berechnungsverfahren zum Erhalten des oben erläuterten Wortauswertemaßes laufen wie folgt:

1. Falls die Länge der Kandidatenzeichen = Länge des Schlüsselworts;
Auswertemaß = (Gesamtsumme der Auswertepunkte für jedes Zeichen)/(Zeichenlänge)
* wenn kein Zeichenkandidat in dem ersten bis vierten Kandidaten vorhanden ist, wird ein Fehlwert (50 HEX) gegeben.

2. Falls die Länge der Kandidatenzeichen = Länge des Schlüsselworts ± 1 ;
Auswertemaß = Summe * 1 (Länge der Kandidatenzeichen, Länge des Schlüsselworts)/MAX (Länge der Kandidatenzeichen, Länge des Schlüsselworts) + Kompensationspunkte (A HEX)

*1: Der Maximalwert, wenn die Auswertepunkte von zwei Richtungen von oben und nach unten addiert werden, um Min. (Länge der Kandidatenzeichen, Länge des Schlüsselworts), und der Gesamtsumme aus der Vorderrichtung werden addiert, indem sie um ein Zeichen nach hinten verschoben werden.

Summe (Länge der Kandidatenzeichen, Länge Schlüsselwort) = Max PS (i)

$$(1 \leq i \leq M + 1)$$

$$PS\ 1(i) = \sum_{n=1}^i S(n) = \text{Gesamtsumme der Auswertepunkte von oben}$$

$$PS\ 2(i) = \sum_{n=i}^M S(n) = \text{Gesamtsumme der Auswertepunkte von unten}$$

wobei

$M = \text{Min}(\text{Lng}(\text{Kandidat}), \text{Lng}(\text{Schlüsselwort}))$

$\text{Lng}(*): *$ Länge der Zeichen des Zugs

$S(n) = \text{Auswertepunkte beim } n\text{-ten Zeichen}$

$$PS(i) = PS\ 1(i-1) + PS\ 2(i); 2 \leq i \leq M$$

$$= PS\ 2(i); i = 1$$

$$= PS\ 1(i); i = M+1$$

Damit wird bei der erfindungsgemäßen Vorrichtung die Gesamtsumme der Auswertepunkte jedes Zeichens der Wortkandidaten dividiert durch die größere Zahl aus entweder der Anzahl der Zeichen im Schlüsselwort oder der Anzahl von Zeichen in dem Wort, um sie zu dem Wortauswertemaß zu machen, so daß es möglich ist, das exakte Wortauswertemaß zu ermitteln und so exakt und stabil Ergebnisse der Wortrecherche zu erhalten.

Die Wortrecherche bei diesem Ausführungsbeispiel ist nicht beschränkt auf die Vorrichtung, die die Trigrammrecherche durchführt, sondern es ist auch möglich, sie in Verbindung mit anderen Vorrichtungen einzusetzen, zum Beispiel in Verbindung mit dem ersten Ausführungsbeispiel.

Im folgenden wird eine fünfte Ausführungsform erläutert.

Bei dem Aufbau nach Fig. 1 ist die Wortregistrierzone des Wortverzeichnisses 13 in der Postleitzahl-Lesevorrichtung dieser Ausführungsform unterteilt in mehrere Gruppen einer Trigrammrecherche-Worttabelle, einer Direktworttabelle, einer Tabelle für numerische Wörter, einer Spezialworttabelle und dergleichen. Der Wortrecherchenabschnitt 14 ist derart aufgebaut, daß er die Anzahl von Zeichen des das Schlüsselwort bildenden Zeichenzugs sowie die Art des Zeichens erkennt und den entsprechenden Typ der Worttabelle innerhalb des Wortverzeichnisses 13 festgelegt auf der Grundlage der Erkennungsergebnisse, um die Wortrecherche in der Worttabelle auf der Grundlage des Schlüsselworts durchzuführen.

Weiterhin besteht in der Trigrammrecherche-Worttabelle eine Gruppe aus Wortgruppen, die die gleichen 3-Zeichen-Züge enthalten, und die gruppiert sind in Abhängigkeit davon, welche Stellung der 3-Zeichen-Zug in dem Wort einnimmt.

Anhand der Fig. 8 wird die Wortrecherche gemäß dieser Ausführungsform erläutert.

Wenn die Zeichenkandidatengruppe und die Wortbegrenzungsinformation als Ergebnis der Erkennung für jedes Zeichen von dem Zeichenerkennungsabschnitt 12 eingegeben sind, extrahiert der Wortrecherchenabschnitt 14 den Zeichencode, der für jedes Zeichen die höheren Auswertepunkte innerhalb dieser Zeichenkandidatengruppe aufweist, um das Schlüsselwort für ein Wort auf der Grundlage der Wortbegrenzungsinformation zu generieren (Schritt 801).

Fig. 9 zeigt ein Beispiel für die Ergebnisse der Zeichenerkennung. Wie aus der Figur ersichtlich ist, handelt es sich bei dem von Zeichenerkennungsabschnitt 12 kommenden Zeichencode um einen von vier Arten: Numerisches Zeichen, großes Alphazeichen, kleines Alphazeichen und Spezialzeichen. Jeder Kandidat für jedes Zeichen besteht aus den Kandidaten eins bis vier. Jeder Zeichenkandidat besitzt Auswertepunkte, die das Ähnlichkeitsmaß mit dem Standardmuster ausdrücken.

Jeder aus diesen Zeichenkandidaten generierte Kandidat gehört zu einem Trigrammwort, welches aus mehr als vier Zeichen des Alphazeichen-Zuges besteht, dem Direktwort, welches aus weniger als drei Zeichen des Alphazeichen-Zuges besteht, dem numerischen Wort, das aus weniger als sechs Ziffern des numerischen Zeichenzugs oder Ordnungszahlen besteht, oder aus einem Spezialwort.

Nachdem das Schlüsselwort generiert ist, erkennt der Wortrecherchenabschnitt die Anzahl von Zeichen und die Art der Zeichen anhand des Schlüsselworts, um zu bestimmen, ob das Schlüsselwort zum Trigrammrecherchen-Wort, zum Direktwort, zum numerischen Wort oder zum Spezialwort gehört (Schritt 802).

Wenn festgestellt wird, daß das Schlüsselwort zu dem Trigrammrecherchen-Wort gehört (Schritt 803), so erfolgt die Trigrammrecherche des Worts anhand der Trigrammrecherche-Worttabelle innerhalb des Wortverzeichnisses 13 (Schritt 804).

Die Trigrammrecherche erfolgt hier folgendermaßen: Zunächst werden aus der Trigrammrecherche-Worttabelle all diejenigen Wortkandidaten herausgezogen, bei denen der 3-Zeichen-Zug im Schlüsselwort und in dem Wort an der gleichen Stelle in jedem Zeichenzug existiert. Weiterhin wird die Position des 3-Zeichen-Zugs in dem Schlüsselwort um ein Zeichen verschoben, um mit jedem Wortkandidaten verglichen zu werden, und es wird die Anzahl von Übereinstimmungen bei jedem Wort gezählt, um das Wort als ersten Kandidaten zu erhalten, bei dem die höhere Anzahl von Übereinstimmungen vorliegt.

Wenn das Schlüsselwort im Schritt 805 als Direktwort ermittelt wird, so erfolgt die Wortrecherche in der Direktwort-Tabelle innerhalb des Wortverzeichnisses 13 (Schritt 806).

Bei diesem Direktwort handelt es sich um einen Alphazeichen-Zug mit weniger als drei Zeichen, so daß die Arten des Zeichenzugs den Wert der Gesamtsumme 27, zur dritten Potenz erhoben, haben. Bei der Recherche des Direktworts wird also untersucht, ob dieser Wort-ID-Code in der Direktworttabelle vorhanden ist oder nicht, und falls vorhanden, wird der Wort-ID-Code als Wortkandidat geliefert.

Wenn es sich bei dem Schlüsselwort um ein numerisches Wort handelt (Schritt 806), so wird dieses numerische Wort in der Numerikwort-Tabelle innerhalb des Wortverzeichnisses 13 recherchiert (Schritt 807).

Weiterhin wird als Ergebnis dieser Recherche nach einem numerischen Wort der Wort-ID direkt erhalten durch Berechnung aus diesem numerischen Zeichencode von [0] bis [9].

Die obige Beschreibung bezieht sich auf die Recherche des Trigrammrecherche-Worts, des Direktworts und des numerischen Zeichenworts.

Andererseits erfolgt die Recherche des Spezialworts, nachdem die Recherche für jedes Wort nach den Schritten 809 bis 812 abgeschlossen ist, da die Häufigkeit des Erscheinens eines Spezialworts gering ist.

Diese Spezialwörter sind in der Spezialworttabelle zusammen mit den Wort-ID-Codes registriert, so daß jeder Zeichencode in absteigender Reihenfolge im ASCII-Code ausgerichtet ist und die Recherche als Binärrecherche durchgeführt wird.

Bei dieser Ausführungsform der Erfindung wird die zu recherchierende Worttabelle innerhalb des Wortverzeichnisses 13 festgelegt nach Maßgabe der Anzahl von Zeichen des das Schlüsselwort bildenden Zeichenzugs und die Art des Zeichens, um eine zu jeder Worttabelle passende Wortrecherche durchzuführen. Damit ist es möglich, die Geschwindigkeit der Wortrecherche beträchtlich heraufzusetzen.

Das Spezialwort, dessen Erscheinungshäufigkeit gering ist, erfolgt kollektiv nach Beendigung der Recherche von häufig auftretenden Wörtern, so daß dadurch, wenn die Verarbeitung innerhalb begrenzter Zeit durchzuführen ist, die Menge nicht abgeschlossener Verarbeitungsvorgänge reduziert wird.

Bei dieser Ausführungsform erfolgt die Recherche des Spezialworts nach der Durchführung der Trigrammrecherche, der Recherche des Direktworts und der Recherche des numerischen Worts. Wenn die Wortarten innerhalb eines Dokuments jedoch nicht ausgewogen sind, kann man jedes Wort irgendeiner Art derart unter-

bringen, daß die durchzuführende Wortrecherche gemäß der jeweiligen Sorte gemäß der Priorität des Worts durchgeführt wird.

Bei dem oben beschriebenen Ausführungsbeispiel wurde die Erkennung des Zeichenmusters erläutert. Die Erfindung ist jedoch nicht auf diese spezielle Zeichenerkennung beschränkt. Anwendbar ist die Erfindung auch auf eine Vorrichtung, die Stimm-Muster erkennt.

Patentansprüche

1. Mustererkennungsvorrichtung, gekennzeichnet durch:
 eine Musterleseeinrichtung (11), die ein Eingabemuster liest;
 eine Mustererkennungseinrichtung (12), die das Eingabemuster erkennt, indem Auswertepunkte berechnet werden, welche den Grad der Ähnlichkeit zwischen dem von der Leseeinrichtung (11) gelesenen Eingabemuster und einem Standardmuster in mehreren Ebenen ausdrücken;
 eine Wortspeichereinrichtung (13) zum Vorab-Speichern von Wörtern; und
 eine Wortrechercheneinrichtung (14) zum Generieren eines Schlüsselworts für die Wortrecherche, indem Vorzugszeichen gesammelt werden, die höhere von der Mustererkennungseinrichtung (12) erhaltene Auswertepunkte aufweisen, und zum Recherchieren von Wörtern in der Wortspeichereinrichtung (13) nach Maßgabe des Schlüsselworts. 10
2. Mustererkennungsvorrichtung nach Anspruch 1, dadurch gekennzeichnet, daß die Wortrechercheneinrichtung (14) aufweist: eine erste Rechercheneinrichtung zum Generieren des Schlüsselworts für die Wortrecherche auf der Grundlage des Ergebnisses von der Mustererkennungseinrichtung (12), zum Vergleichen des Schlüsselworts mit den in der Wortspeichereinrichtung (13) gespeicherten Wörtern in jeder eine gewisse Anzahl von Zeichen umfassenden Zeichenkette und zum bevorzugten Ausgeben des Worts mit dem höheren Verhältnis der durch Vergleich ermittelten Übereinstimmungszahl, eine zweite Wortrechercheneinrichtung für die Gesamtsumme der Auswertepunkte jedes Zeichens des durch die erste Rechercheneinrichtung recherchierten Worts, gebildet durch die höhere Zahl entweder der recherchierten Wortzahlen oder Schlüsselwortzeichen-Zahlen und zur bevorzugten Ausgabe des ermittelten höheren Verhältnisses als endgültiges Wortrecherchenergebnis, und eine dritte Rechercheneinrichtung, welche die Gesamtsumme der Auswertepunkte jedes Zeichens des durch die zweite Rechercheneinrichtung recherchierten Worts dividiert durch die höhere Zahl von entweder den recherchierten Wortzahlen oder den Schlüsselwortzeichen-Zahlen und zur bevorzugten Ausgabe des erhaltenen höheren Verhältnisses als endgültiges Wortrecherchenergebnis. 15 20 25 30
3. Mustererkennungsvorrichtung, gekennzeichnet durch:
 eine Musterleseeinrichtung (11), die ein Eingabemuster liest;
 eine Mustererkennungseinrichtung (12), die das durch die Musterleseeinrichtung gelesene Eingabemuster erkennt;
 eine Wortspeichereinrichtung (13), in der vorab Wörter gespeichert werden; und
 eine Wortrechercheneinrichtung (14), die ein Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung erhaltenen Ergebnisse generiert, das Schlüsselwort vergleicht mit den in der Wortspeichereinrichtung (13) gespeicherten Wörtern für jede gewisse Anzahl von Zeichen einer Zeichenkette, und die bevorzugt das Wort mit dem höheren Verhältnis der durch Vergleich gebildeten Übereinstimmungszahl als endgültiges Wortrecherchenergebnis ausgibt. 35 40
4. Mustererkennungsvorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß die Mustererkennungsvorrichtung (12) das Eingabemuster dadurch erkennt, daß sie Auswertepunkte berechnet, welche das Ähnlichkeitsmaß zwischen dem von der Musterleseeinrichtung gelesenen Eingabemuster und einem Standardmuster ausdrücken, und daß die Wortrechercheneinrichtung aufweist: eine erste Rechercheneinrichtung, die das Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung erhaltenen Ergebnisse generiert, das Schlüsselwort vergleicht mit den in der Wortspeichereinrichtung gespeicherten Wörtern für jede gewisse Anzahl von Zeichen einer Zeichenkette, und vorzugsweise das Wort mit dem höheren Verhältnis von durch Vergleich gewonnener Übereinstimmungszahl ausgibt, und eine zweite Wortrechercheneinrichtung, die jedes von der ersten Rechercheneinrichtung kommende Wort mit dem Schlüsselwort — in jede Wort vergleicht und bevorzugt als endgültiges Ergebnis der Wortrecherche dasjenige Wort ausgibt, welches die höhere Anzahl übereinstimmender Zeichen aufweist. 45 50
5. Mustererkennungsvorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß die Wortrechercheneinrichtung aufweist: eine erste Rechercheneinrichtung, die das Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung erhaltenen Ergebnisse generiert, das Schlüsselwort vergleicht mit den in der Wortspeichereinrichtung gespeicherten Wörtern für jede gewisse Anzahl von Zeichen einer Zeichenkette, und bevorzugt das Wort mit dem höheren Verhältnis von durch Vergleich gewonnener Übereinstimmungszahl ausgibt, und eine zweite Wortrechercheneinrichtung, die die Gesamtsumme der Auswertepunkte für jedes Zeichen des von der ersten Rechercheneinrichtung recherchierten Worts dividiert durch die höhere Anzahl von entweder der recherchierten Wortzahlen oder der Schlüsselwortzeichen-Anzahlen, und die bevorzugt als letztlches Wortrecherchen-Ergebnis das höhere Verhältnis ausgibt. 55 60
6. Mustererkennungsvorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß die Mustererkennungseinrichtung das Eingabemuster erkennt durch Berechnen von Auswertepunkten, welche das Ähnlichkeitsmaß zwischen dem von der Musterleseeinrichtung (11) gelesenen Eingabemuster und einem Standardmuster ausdrücken, und daß die Wortrechercheneinrichtung (13) enthält: eine erste Rechercheneinrichtung, die das Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung erhalte- 65

nen Ergebnisse generiert, das Schlüsselwort vergleicht mit den in der Wortspeichereinrichtung gespeicherten Wörtern für jeweils eine gewisse Anzahl von Zeichen einer Zeichenkette, und die bevorzugt das Wort mit dem höheren Verhältnis der durch Vergleich gewonnenen Übereinstimmungszahl ausgibt, eine zweite Wortrechercheneinrichtung, die jedes Zeichen des von der ersten Rechercheneinrichtung recherchierten Worts mit dem Schlüsselwort vergleicht, um bevorzugt dasjenige Wort auszugeben, welches das höhere Verhältnis von durch Vergleich ermittelten Übereinstimmungszahlen aufweist, und daß eine dritte Rechercheneinrichtung vorgesehen ist, welche die Gesamtsumme der Auswertepunkte für jedes Zeichen des von der zweiten Rechercheneinrichtung recherchierten Worts dividiert durch die größere Anzahl von entweder den Anzahlen recherchiert Wörter oder den Anzahlen von Schlüsselwortzeichen, und die bevorzugt das höhere Verhältnis als endgültiges Ergebnis der Wortrecherche ausgibt.

7. Zeichenerkennungsvorrichtung, gekennzeichnet durch:

eine Musterleseeinrichtung (11) zum Lesen von Eingabemustern;

eine Mustererkennungseinrichtung (12) zum Erkennen des von der Musterleseeinrichtung (11) gelesenen Eingabemusters;

eine Wortspeichereinrichtung (13) zum Vorab-Speichern von Wörtern;

eine erste Wortrechercheneinrichtung (13), die ein Schlüsselwort für die Wortrecherche auf der Grundlage der von der Zeichenerkennungseinrichtung (12) erhaltenen Ergebnisse generiert und in der Wortspeichereinrichtung auf der Grundlage des Schlüsselworts recherchiert; und

eine zweite Wortrechercheneinrichtung (16), die jedes von der ersten Wortrechercheneinrichtung ausgegebene Wort mit dem Schlüsselwort in jedem Zeichen vergleicht und bevorzugt als letztlisches Ergebnis der Wortrecherche das Wort ausgibt, welches die meisten übereinstimmenden Zeichen aufweist.

8. Mustererkennungsvorrichtung, gekennzeichnet durch:

eine Musterleseeinrichtung (11) zum Lesen von Eingabemustern;

eine Worterkennungseinrichtung (12), die das Eingabemuster erkennt, indem sie Auswertepunkte berechnet, welche das Ähnlichkeitsmaß zwischen dem von der Musterleseeinrichtung gelesenen Eingabemuster und einem Standardmuster ausdrücken;

eine Wortspeichereinrichtung für die Vorab-Speicherung von Wörtern;

eine erste Wortrechercheneinrichtung (13), die ein Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung (12) erhaltenen Ergebnisse generiert und in der Wortspeichereinrichtung (13) auf der Grundlage des Schlüsselworts recherchiert; und

eine zweite Wortrechercheneinrichtung, welche die Gesamtsumme der Auswertepunkte jedes von der ersten Wortrechercheneinrichtung recherchierten Worts dividiert durch die höhere Zahl von entweder der Anzahl recherchiert Wörter oder der Anzahl von Schlüsselwort-Zeichen, und die als endgültiges Wortrecherchen-Ergebnis bevorzugt das höhere Verhältnis ausgibt.

9. Mustererkennungsvorrichtung, gekennzeichnet durch:

eine Musterleseeinrichtung (11) zum Lesen von Eingabemustern;

eine Mustererkennungseinrichtung (12) zum Erkennen von von der Musterleseeinrichtung (11) gelesenen Eingabemustern;

eine Wortspeichereinrichtung (13), die vorab nach gewissen Gruppen-Sorten klassifizierte Wörter speichert;

eine Wortrechercheneinrichtung (14), die ein Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung (12) erhaltenen Ergebnisse generiert, um die Art von Zeichenketten zu identifizieren, die das Schlüsselwort bilden, und die die Wörter nach Maßgabe von Wortgruppen innerhalb der Wortspeichereinrichtung auf der Grundlage des identifizierten Ergebnisses recherchiert.

10. Mustererkennungsvorrichtung nach Anspruch 9, dadurch gekennzeichnet, daß die Wortrechercheneinrichtung das Schlüsselwort für die Wortrecherche auf der Grundlage der von der Mustererkennungseinrichtung erhaltenen Ergebnisse generiert, die Art von das Schlüsselwort bildenden Zeichenketten identifiziert und die Wörter nach Maßgabe von Wortgruppen innerhalb der Wortspeichereinrichtung auf der Grundlage des identifizierten Ergebnisses recherchiert, um einer für jede Art von Wörtern vorab definierten Priorität zu folgen.

Hierzu 8 Seite(n) Zeichnungen

FIG. 1

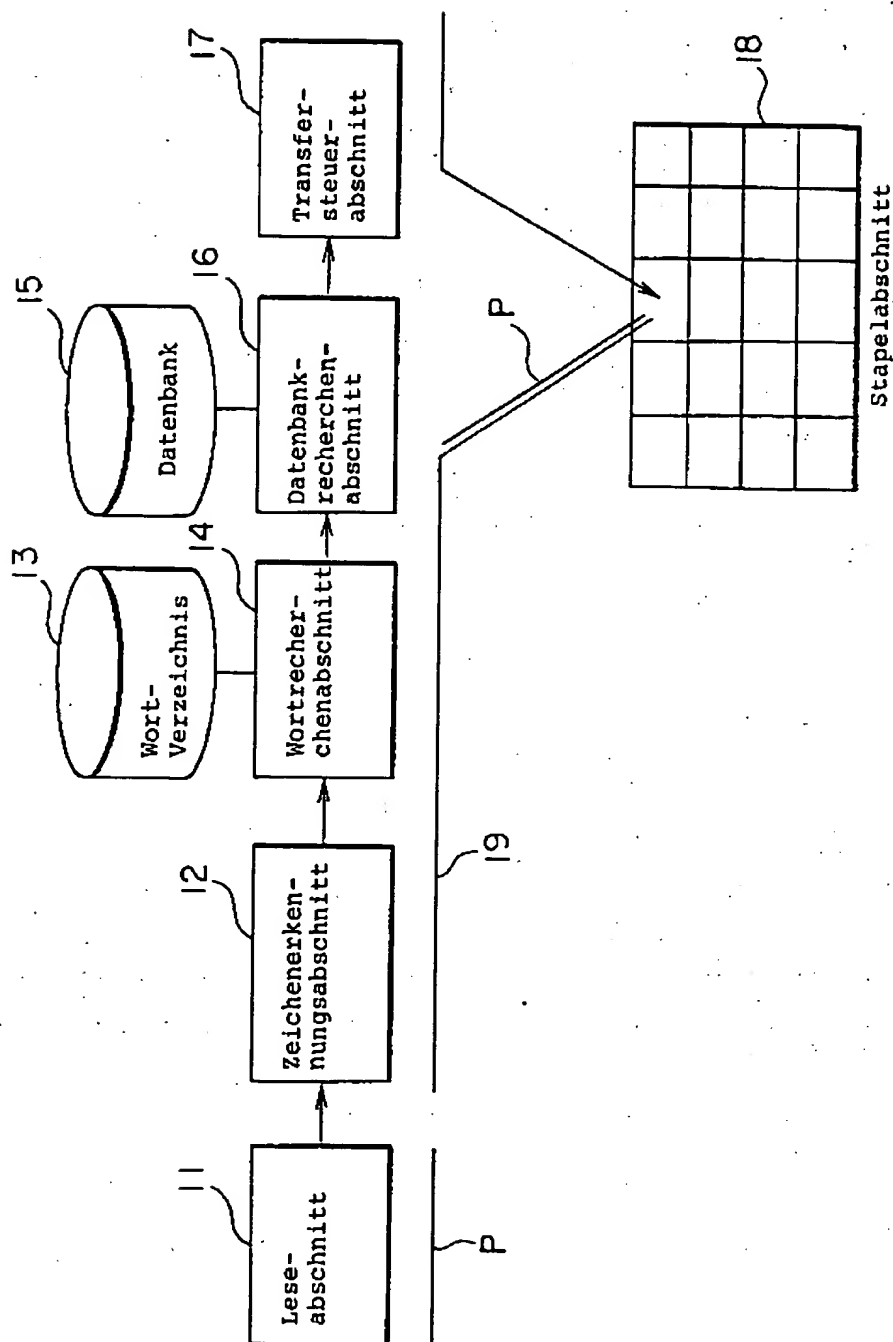


FIG. 2

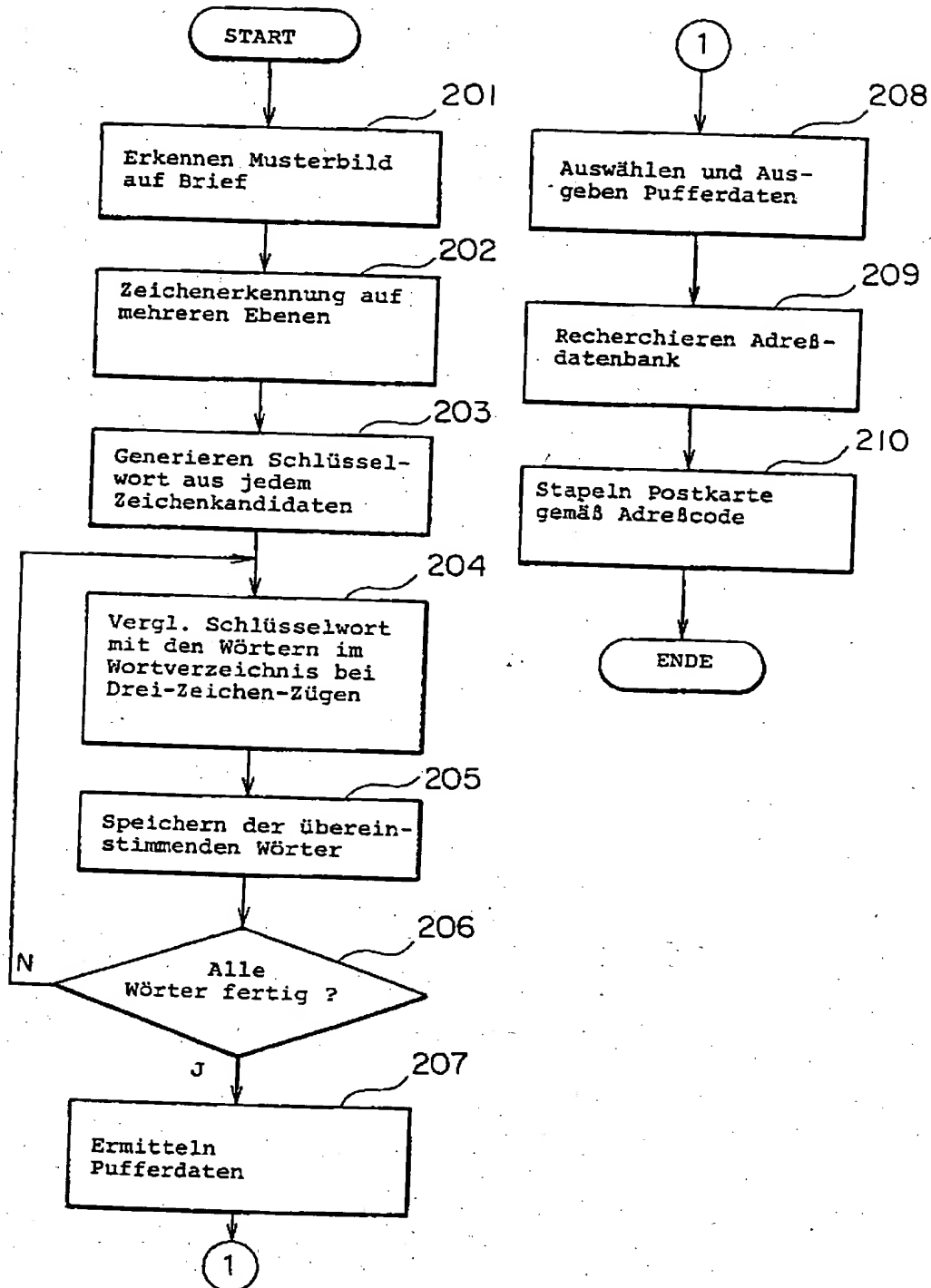


FIG. 3

	erster Kandidat	zweiter Kandidat	dritter Kandidat	vierter Kandidat	
Muster 1	R 90	P 79	F 78	A 6e	Zeichenkandidat für erstes Zeichen
Muster 2	P 86	R 84	F 69	A 67	
Muster 1	O 98	C 86	G 71	E 6d	Zeichenkandidat für zweites Zeichen
Muster 2	O 88	C 78	D 68	G 58	
Muster 1	O 98	C 86	G 71	E 6d	Zeichenkandidat für drittes Zeichen
Muster 2	O 88	C 78	D 68	G 58	
Muster 1	N 98	M 78	H 68	I 58	Zeichenkandidat für viertes Zeichen
Muster 2	M 88	N 82	H 66	W 62	

FIG. 4

erster Kandidat	zweiter Kandidat	dritter Kandidat	vierter Kandidat	
R 90	P 86	F 78	A 6e	Zeichenkandidat für erstes Zeichen
O 98	C 88	G 71	E 6d	Zeichenkandidat für zweites Zeichen
O 98	C 88	G 71	E 6d	Zeichenkandidat für drittes Zeichen
N 98	M 88	H 68	W 62	Zeichenkandidat für viertes Zeichen

FIG. 5

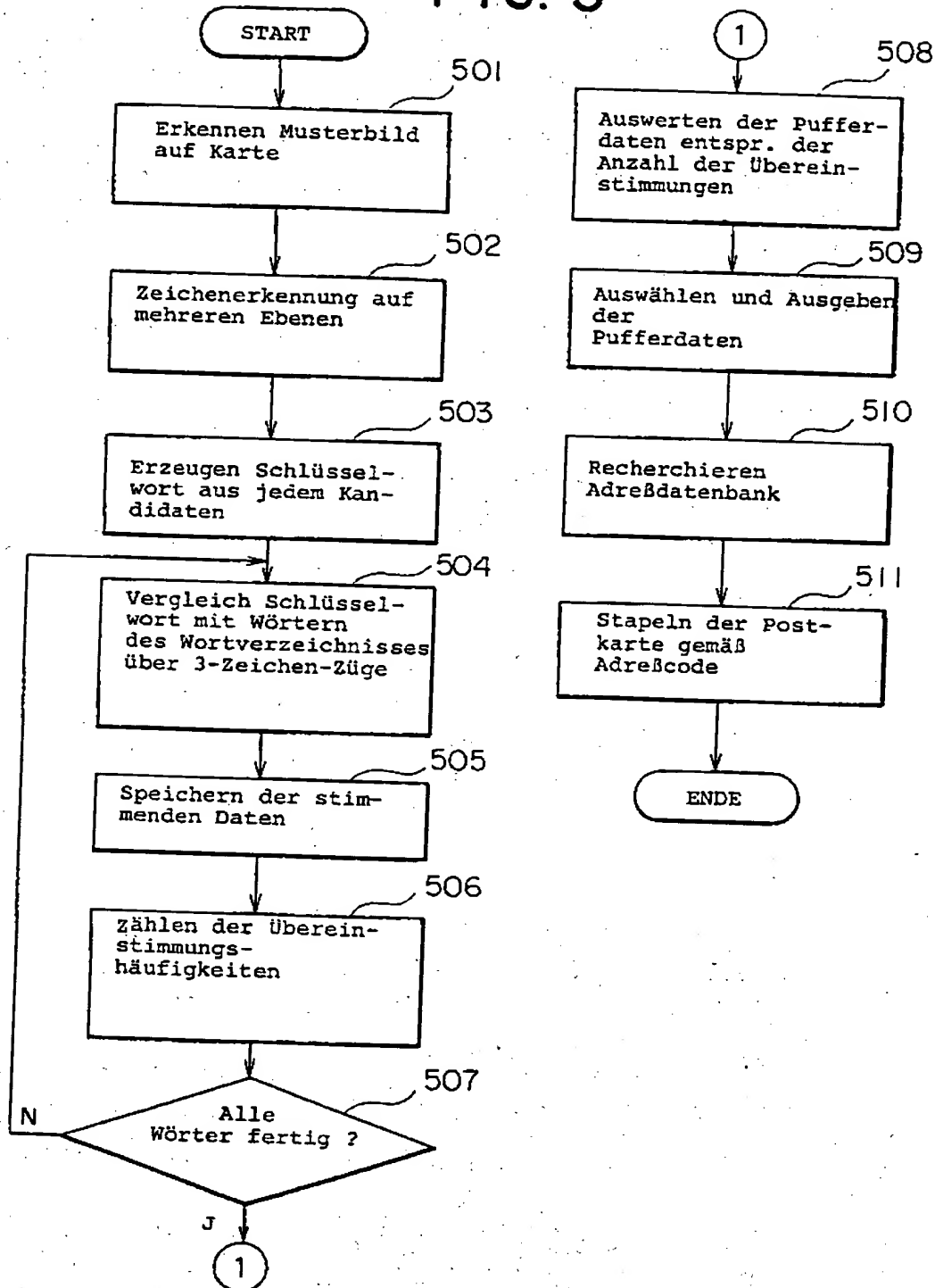


FIG. 6

Zählpuffer

Trigramm-Zählpuffer

0000 0000 0000 0000 0000 0000 0000 0000
 0000 0000 0001 0000 0005 002c 00e3 00ff

Wortzahl bei Wortzahl bei Wortzahl bei
 6-maliger 4-maliger 2-maliger
 Übereinstim- Übereinstim- Übereinstimmung
 mung mung

Datenpuffer

TRIGRAMM-ANSWER-PUFFER

0400a800 (Datenpuffer für 6-malige Übereinst.)
 ANGLES (50014e71)

0400b000 (Datenpuffer für 4-malige Übereinst.)
 ANGELES (6002e849) ANGLERS (6001c775)
 BANGLES (6007d881) MANGLES (60052976)
 PANGLES (60063f48)

0400b400 (Datenpuffer für 3-malige Übereinst.)
 ANGES (40067f31) ANGLEM (50062af0)
 ANGLE Y (50067066) ANGLE (4006b4e3)
 ANGLASS (6002da72) ANGLEA (5005e57c)
 ANGIES (5001e6c8) ANGLER (50021d64)
 ANGLOS (5001979f) TANTLER (60078958)
 ZANGLE (5006dee3) BANGLE (50082ee2)
 LANGLE (50087818) LANGLEY (600913bc)
 WANGLE (60089f2a) RANGLEY (600140d7)
 LANGLE (60014fae) DANGLE (60051484)
 TANGLE (5004b340) NANGLE (5002879f)
 KANGLEY (60020d72) TANGLED (6001ffc2)
 RANGLE (5000a539) TANGLED (60054ee0)
 PANGLE (500695ad) TANGLEY (600a3cd0)
 INGLESE (6001a81c) ENGLES (50096a81)
 INGLES (60022125) ENGLES (60038a86)
 INGLES (50078816) SINGLES (6000d9b0)
 MINLES (6008ac9b) DINGLES (60056b29)
 TINGLES (6007dffa) SYNGLES (6007dffb)
 JUNGLES (6007dc3d) GINGLES (600680fd)
 FINGLES (60097938) SINGLE (5001301d)
 SAGLE (4003c349) ANTLES (500a0485)
 ANKLES (50089975) ANTILES (6000d4b0)

FIG. 7

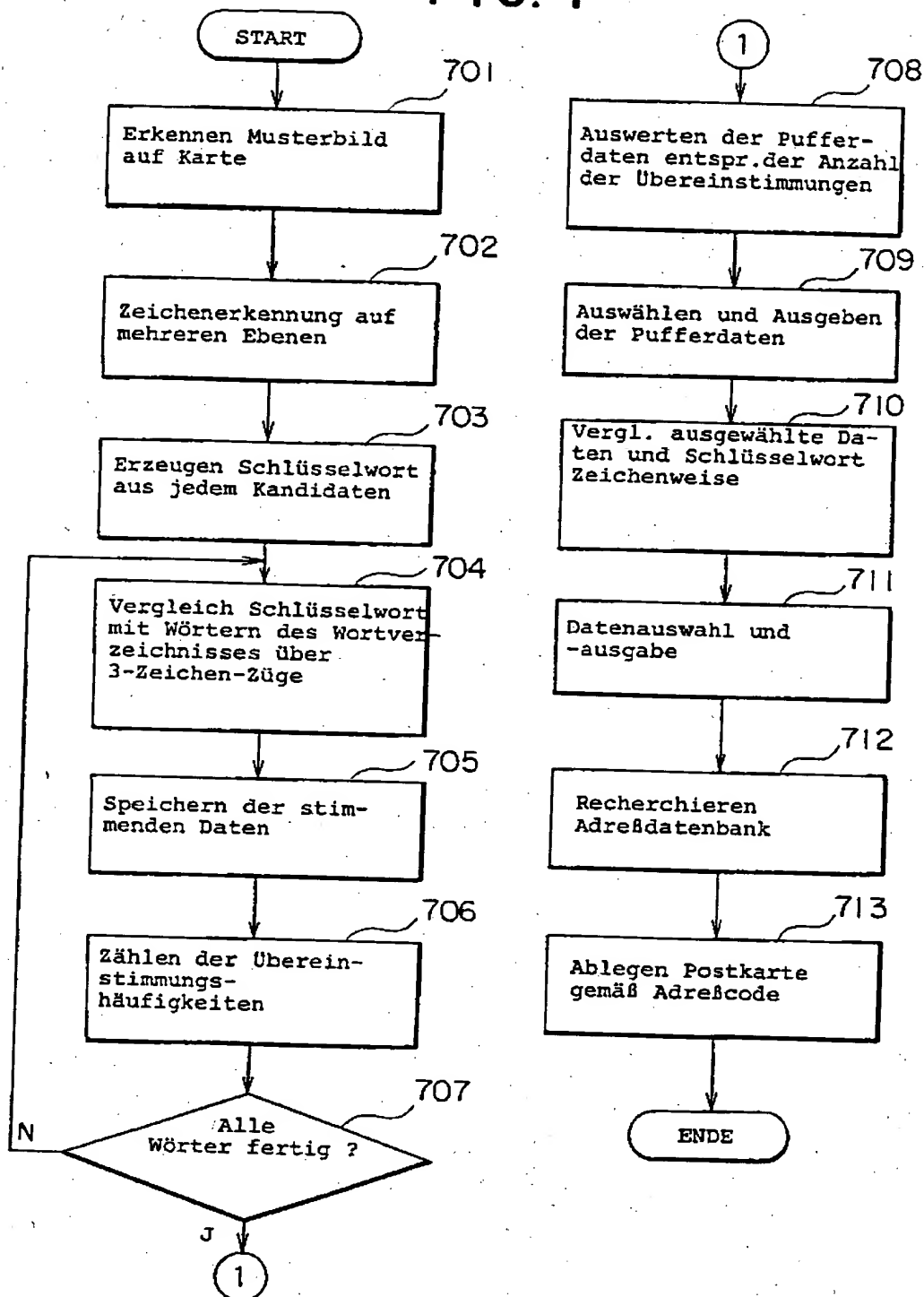


FIG. 8

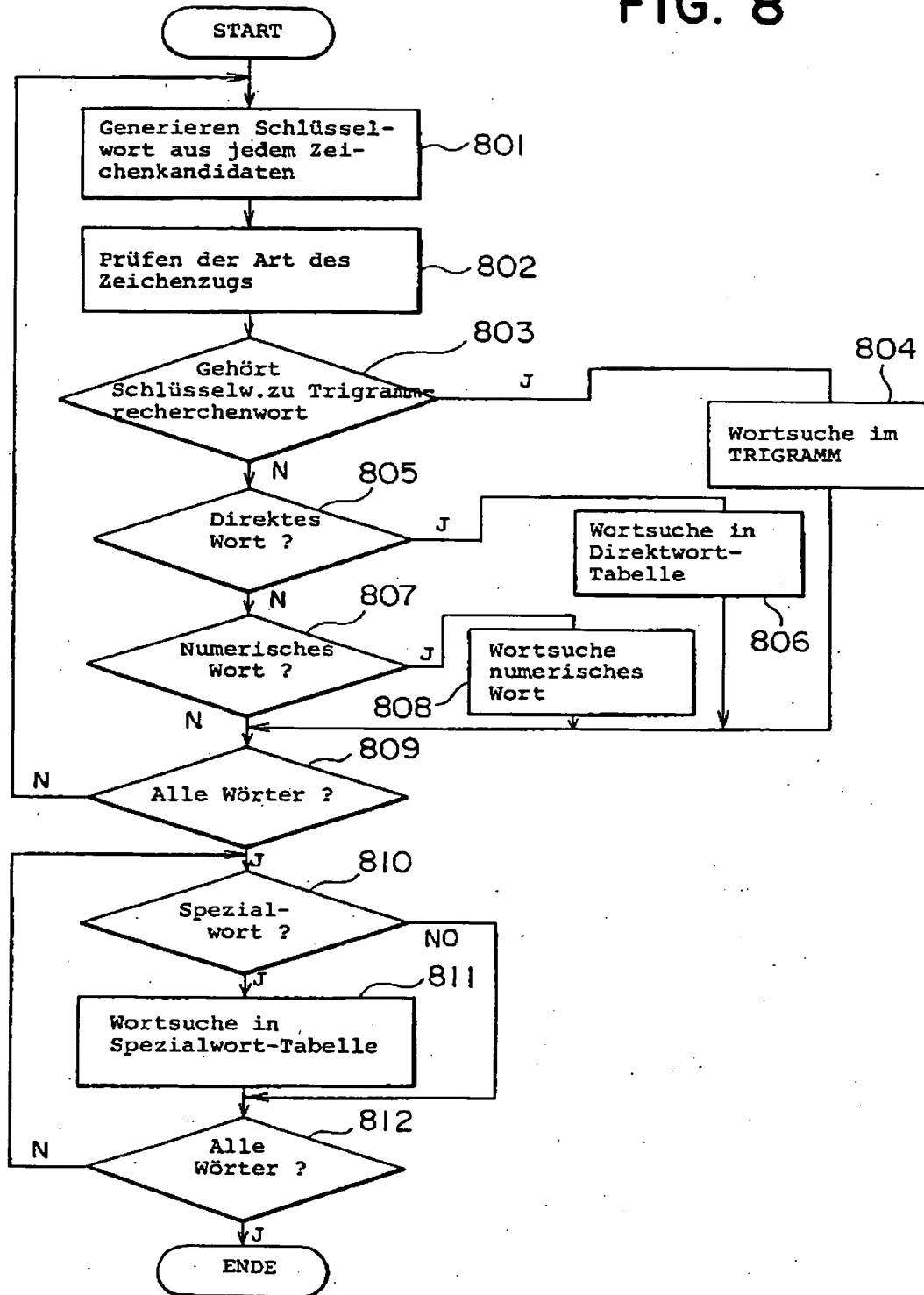


FIG. 9

	erster Kandidat	zweiter Kandidat	dritter Kandidat	vierter Kandidat	
	Zeichencode-Auswertepunkt				
erstes Zeichen	000a 3584	000c 3682	0010 3870	0006 3365	— 5 6 8 3
	0125 5398	0103 4278	0108 4568	0000 0058	— S E B
	0225 7398	020d 6778	0220 7168	0203 6258	— s g q b
	0304 2484	0320 2664	0317 2c54	0326 5d50	— \$ 8 . i
	— numerisches Zeichen				
	— engl. Großbuchstaben				
	— engl. Kleinbuchstaben				
	— Spezialwort				
zweites Zeichen	0001 3084	0013 3964	0009 3454	000d 3650	0 9 4 6
	0129 5598	011d 4176	0107 4468	0000 0058	U O D
	0228 7598	0217 6c78	022a 7668	021c 6158	u i v o
	0326 5d60	030a 2858	0315 2740	0331 7d30	} (,)
drittes Zeichen	0003 3198	0000 3078	000f 3768	0009 3458	1 0 7 4
	0110 4984	0116 4c64	0121 5854	0127 5450	I L X T
	0217 6c98	0211 6978	0227 7468	020a 6658	i l i f
	030d 2984	0326 5d74	030b 2864	0330 7d54) j (}
viertes Zeichen	0004 3284	0010 3882	000c 3670	000b 3665	2 8 6 5
	0109 4560	0124 5340	0104 4330	0103 4220	E S C B
	0208 6598	0204 6378	0225 7368	020d 6758	e c s g
	0301 4084	0317 2c64	030b 2854	0320 2650	Ⓢ (8
fünftes Zeichen	0004 3260	0010 3858	000f 3740	0006 3330	2 8 7 3
	010a 4660	0116 4c40	0119 4d30	0110 4920	F L M I
	0227 7498	020e 6878	0205 6368	021a 6a58	t h c n
	0329 5b84	031d 3b64	0326 5d54	0320 2650	[;] B